



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :  C07B 59/00, C12N 1/20, 5/00, 5/06, C12P 21/00, G01N 24/08, C07K 14/59		A1	(11) International Publication Number: <b>WO 99/11589</b>  (43) International Publication Date: 11 March 1999 (11.03.99)
(21) International Application Number: PCT/US98/18197  (22) International Filing Date: 2 September 1998 (02.09.98)		Hall Street, Kettlebridge, Fife KY7 7QF (GB). ZIMMER-MAN-WILSON, Carrie, Ann [US/US]; 2642 Westchester Avenue, Ellicott City, MD 21043 (US). OLIVER, Eddie, James, Jr. [US/US]; Apartment 204, 7607 Eastern Avenue, Takoma Park, MD 20912 (US).	
(30) Priority Data:  08/921,554 2 September 1997 (02.09.97) US		(74) Agents: SULLIVAN, Michael, G. et al.; Rothwell, Figg, Ernst & Kurz, Suite 701 East, Columbia Square, 555 13th Street, N.W., Washington, DC 20004 (US).	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application  US Filed on Not furnished (CIP) Not furnished		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71) Applicant (for all designated States except US): MARTEK BIOSCIENCES CORPORATION [US/US]; 6480 Dobbin Road, Columbia, MD 21045 (US).		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(72) Inventors; and			
(75) Inventors/Applicants (for US only): ANDERSON, Frank, Elbert, III [US/US]; Apartment 163, 17022 King James Way, Gaithersburg, MD 20877 (US). BROWN, Jonathan, Miles [GB/US]; 13912 North Gate Drive, Silver Spring, MD 20906 (US). COUGHLIN, Philip, Edward [GB/US]; 924 Wilson Road, Wilmington, DE 19803 (US). HOMANS, Steven, William [GB/GB]; 4 Bucklerscroft, Kellas, Dundee DDS 3PQ (GB). WELLER, Charles, Tobias [GB/GB]; 6			
(54) Title: 13C, 15N, 2H LABELED PROTEINS FOR NMR STRUCTURE DETERMINATIONS AND THEIR PREPARATION			
(57) Abstract			
<p>This invention is concerned with determining the three-dimensional structure of biological macromolecules such as proteins. In particular, it is concerned with methods for rapidly determining protein structure by NMR, by providing methods for simplifying NMR spectra using labeled proteins prepared from specifically isotopically labeled amino acids having at least two isotopes of 13C, 15N and 2H in-the-backbone, and methods for making these labeled proteins, e.g., by cultivation of a microbial culture containing said labeled amino acids.</p>			

## 13C,15N,2H LABELED PROTEINS FOR NMR STRUCTURE DETERMINATIONS AND THEIR PREPARATION

### Field of the Invention

This invention is concerned with determining the  
5 three-dimensional structure of biological  
macromolecules, especially proteins. In particular, it  
is concerned with methods for rapidly determining  
protein structures by NMR spectroscopy, by providing  
methods for simplifying NMR spectra using labeled  
10 proteins prepared from specifically isotopically  
labeled amino acids, and the means whereby these  
labeled proteins and amino acids may be obtained.

### Background of the Invention

For many years, there has been intense interest in  
15 determining the three-dimensional structures of  
biological macromolecules, particularly proteins. So  
called "structure-function" studies have been carried  
out to determine the structural features of a molecule,  
or class of molecules, that are important for  
20 biological activity. Since the pioneering work of  
Perutz and coworkers on the structure of hemoglobin  
(Perutz, M.F. et al., *Nature*, 185:416-22 (1960)) and  
that of Watson and Crick on DNA in the 1950's (Watson,

J.D. and Crick, F.H.C., *Nature*, 171:737 (1953), both of which led to the respective scientists receiving the Nobel Prize, this field has been of major importance in the biological sciences.

5 More recently, the concept of "rational drug design" has evolved. This strategy for the design of drugs involves determining the three-dimensional structure of an "active part" of a particular biological molecule, such as a protein. Knowing the  
10 three-dimensional structure of the active part can enable scientists to design a synthetic analogue of the active part that will block, mimic or enhance the natural biological activity of the molecule. (Appelt, K. et al., *J. Med. Chem.*, 34:1925 (1991)). The  
15 biological molecule may, for example, be a receptor, an enzyme, a hormone, or other biologically active molecule. Determining the three-dimensional structures of biological molecules is, therefore, of great practical and commercial significance.

20 The first technique developed to determine three-dimensional structures was X-ray crystallography. The structures of hemoglobin and DNA were determined using this technique. In X-ray crystallography, a crystal (or fiber) of the material to be examined is bombarded  
25 with a beam of X-rays which are refracted by the atoms of the ordered molecules in the crystal. The scattered X-rays are captured on a photographic plate which is then developed using standard techniques. The diffracted X-rays are thus visualized as a series of  
30 spots on the plate and from this pattern, the structure of the molecules in the crystal can be determined. For larger molecules, it is frequently necessary to

crystallize the material with a heavy ion, such as ruthenium, in order to remove ambiguity due to phase differences.

More recently, a second technique, nuclear magnetic resonance (NMR) spectroscopy, has been developed to determine the three-dimensional structures of biological molecules, particularly proteins. NMR was originally developed in the 1950's and has evolved into a powerful procedure to analyze the structure of small compounds such as those with a molecular weight of  $\leq$  1000 Daltons. Briefly, the technique involves placing the material to be examined (usually in a suitable solvent) in a powerful magnetic field and irradiating it with radio frequency (rf) electromagnetic radiation. The nuclei of the various atoms will align themselves with the magnetic field until energized by the rf radiation. They then absorb this resonant energy and re-radiate it at a frequency dependent on i) the type of nucleus and ii) its atomic environment. Moreover, resonant energy can be passed from one nucleus to another, either through bonds or through three-dimensional space, thus giving information about the environment of a particular nucleus and nuclei in its vicinity.

However, it is important to recognize that not all nuclei are NMR active. Indeed, not all isotopes of the same element are active. For example, whereas "ordinary" hydrogen,  $^1\text{H}$ , is NMR active, heavy hydrogen (deuterium),  $^2\text{H}$ , is not active in the same way. Thus, any material that normally contains  $^1\text{H}$  hydrogen can be rendered "invisible" in the hydrogen NMR spectrum by replacing all the  $^1\text{H}$  hydrogens with  $^2\text{H}$ . It is for this

reason that NMR spectroscopic analyses of water-soluble materials frequently are performed in  $^2\text{H}_2\text{O}$  to eliminate the water signal.

Conversely, "ordinary" carbon,  $^{12}\text{C}$ , is NMR inactive whereas the stable isotope,  $^{13}\text{C}$ , present to about 1% of total carbon in nature, is active. Similarly, while "ordinary" nitrogen,  $^{14}\text{N}$ , is nmr active, it has undesirable properties for NMR and resonates at a different frequency from the stable isotope  $^{15}\text{N}$ , present to about 0.4% of total nitrogen in nature. For small molecules, these low level natural abundances were sufficient to generate the required experimental information, provided that the experiment was conducted with sufficient quantities of material and for a sufficient time.

As advances in hardware and software were made, the size of molecules that could be analyzed by these techniques increased to about 10kD, the size of a small protein. Thus, the application of NMR spectroscopy to protein structural determinations began only a few years ago. It was quickly realized that this size limit could be raised by substituting the NMR inactive isotopes  $^{14}\text{N}$  and  $^{12}\text{C}$  in the protein with the NMR active stable isotopes  $^{15}\text{N}$  and  $^{13}\text{C}$ .

Over the past few years, labeling proteins with  $^{15}\text{N}$  and  $^{15}\text{N}/^{13}\text{C}$  has raised the analytical molecular size limit to approximately 15kD and 40kD, respectively. More recently, partial deuteration of the protein in addition to  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeling has increased the size of proteins and protein complexes still further, to approximately 60-70kD. See Shan et al., *J. Am.*

*Chem. Soc.*, 118:6570-6579 (1996) and references cited therein.

Isotopic substitution is usually accomplished by growing a bacterium or yeast, transformed by genetic engineering to produce the protein of choice, in a growth medium containing  $^{13}\text{C}$ -,  $^{15}\text{N}$ - and/or  $^2\text{H}$ -labeled substrates. In practice, bacterial growth media usually consist of  $^{13}\text{C}$ -labeled glucose and/or  $^{15}\text{N}$ -labeled ammonium salts dissolved in  $\text{D}_2\text{O}$  where necessary. Kay, L. et al., *Science*, 249:411 (1990) and references therein and Bax, A., *J. Am. Chem. Soc.*, 115, 4369 (1993). More recently, isotopically labeled media especially adapted for the labeling of bacterially produced macromolecules have been described. See U.S. patent 5,324,658.

The goal of these methods has been to achieve universal and/or random isotopic enrichment of all of the amino acids of the protein. By contrast, some workers have described methods whereby certain residues can be relatively enriched in  $^1\text{H}$ ,  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ . For example, Kay et al., *J. Mol. Biol.*, 263, 627-636 (1996) and Kay et al., *J. Am. Chem. Soc.*, 119, 7599-7600 (1997) have described methods whereby isoleucine, alanine, valine and leucine residues in a protein may be labeled with  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ , but specifically labeled with  $^1\text{H}$  at the terminal methyl position. In this way, study of the proton-proton interactions between some of the hydrophobic amino acids may be facilitated. Similarly, a cell-free system has been described by Yokoyama et al., *J. Biomol. NMR*, 6(2), 129-134 (1995)., wherein a transcription-translation system derived from

*E. coli* was used to express human Ha-Ras protein incorporating  $^{15}\text{N}$  serine and/or aspartic acid.

These methods are important, in that they provide additional means for interpreting the complex spectra obtained from proteins. However, it should be noted that the Kay et al. methods are limited to the aliphatic amino acids described above. By contrast, the method described by Yokoyama will facilitate the selective enrichment of any amino acid, but is limited to those proteins that can be expressed in a cell-free system. Glycoproteins, for example, may not be expressed in this system.

Techniques for producing isotopically labeled proteins and macromolecules, such as glycoproteins, in mammalian or insect cells have been described. See U.S. patents 5,393,669 and 5,627,044; Weller, C.T., *Biochem.*, 35, 8815-23 (1996) and Lustbader, J.W., *J.Biomol. NMR*, 7, 295-304 (1996). Weller et al. applied these techniques to the determination of the structure of a glycoprotein including its glycosyl sidechain.

While the above techniques represent remarkable advances in this field, they each suffer from certain disadvantages. For example, all are time-consuming. In X-ray crystallographic methods, crystals can take years to form before the experiment even starts. In NMR spectroscopy, although the protein sample may be used immediately in the NMR experiment, processing the data obtained, i.e., analyzing which signal comes from which set of which atoms (the "assignments"), may also take years. Modern drug discovery research depends heavily on knowledge of the structures of biologically

active macromolecules. This research would benefit substantially from enhancements in the capabilities and speed of three-dimensional structural analyses of proteins and other macromolecules.

5 In the past few years, growth in discovering alternative, rapid methods for the identification of candidate drugs has occurred. Genomic techniques, using rapid DNA sequencing methods and computer assisted homology identification, have enabled the 10 rapid identification of target proteins as potential drug candidates. O'Brien, C., *Nature*, 385 (6616):472 (1997). Once identified, a target protein can be quickly produced using modern recombinant technology. Combinatorial chemistry, wherein large numbers of 15 chemical compounds are simultaneously synthesized on plastic plates, frequently by robots, has revolutionized the synthesis of drug candidates, with tens of thousands of compounds ("libraries") able to be synthesized in a few months. See Gordon, F.M. et al., 20 *J. Mol. Chem.*, 37(10), 1385-1401 (1994). The library is then "screened" by allowing each member of the library to come into contact with the target protein. Those that bind are identified, and similar compounds 25 are synthesized and screened. The whole process continues in an iterative manner until a drug candidate of suitably high binding affinity has been identified. One variation of this screening strategy has recently been published by Fesik et al., *Science*, 274, 1531-34 (1996), wherein the screening of the libraries takes 30 place using NMR against an isotopically labeled protein and the binding is detected from perturbations in the NMR spectrum.

Prior knowledge of the three-dimensional structure of a target protein can enable the design of a "focused" combinatorial library, thereby increasing the likelihood of finding potential drug candidates that 5 interact with the biological molecule of interest. However, whereas genomic and combinatorial chemistry each can be performed in months, known methods for protein structural determinations usually take much longer. Therefore, there is a need for methods to 10 increase the speed with which high resolution structures of proteins, including those that are the targets of potential drug candidates, may be determined.

Summary of the Invention

15 The present invention provides novel labeled proteins that are isotopically labeled in the backbone structure, but not in the amino acid side chains. The invention also provides novel cell culture media that contain one or more amino acids isotopically labeled in 20 the backbone structure but not in the side chain, and methods for making a labeled protein by cultivating a protein-producing cell culture on such a culture medium.

In another aspect, the invention provides a method 25 for determining the three-dimensional structure of a protein wherein at least one of the amino acids in the protein is specifically labeled in its backbone but not its side chain with any combination of the NMR isotopes  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ .

In yet another aspect of the present invention, a method is provided for rapidly assigning the signals in the NMR spectrum of a protein wherein at least one of the amino acids in the protein is specifically labeled 5 in its backbone, but not its side chain with any combination of the NMR isotopes  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ .

In preferred embodiments of these various aspects of the invention, the amino acids contained in the culture media and incorporated into the protein 10 structure are labeled in the backbone with  $^{13}\text{C}$  and  $^{15}\text{N}$  and optionally with  $^2\text{H}$ .

#### Brief Description of the Drawings

Figures 1A and 1B show the HNCA spectra, obtained 15 with deuterium decoupling, of N-acetylated,  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  backbone labeled and 50% deuterated, fully  $^{13}\text{C}/^{15}\text{N}$  labeled phenylalanine, respectively.

Figures 2A and 2B show the HNCA spectra, obtained 20 with deuterium decoupling, of N-acetylated, backbone labeled and 50% deuterated, fully  $^{13}\text{C}/^{15}\text{N}$  labeled phenylalanine, respectively, dissolved in fully deuterated glycerol- $\text{H}_2\text{O}$  85:15 v/v, acquired at 0° C.

Figure 3 shows the HNCA spectrum, obtained without 25 deuterium decoupling, of HCG  $\beta$ -subunit, in which the valine, leucine and phenylalanine are triple backbone labeled with  $^2\text{H}$  (50%),  $^{13}\text{C}$  and  $^{15}\text{N}$ .

Figure 4 shows the HNCA spectrum, obtained with deuterium decoupling, of HCG  $\beta$ -subunit, in which the valine, leucine and phenylalanine are triple backbone 30 labeled with  $^2\text{H}$  (50%),  $^{13}\text{C}$  and  $^{15}\text{N}$ .

Detailed Description of the Invention

The invention provides a means for rapidly determining the three-dimensional structure of proteins by NMR. As described in further detail below, this improvement in NMR spectroscopic techniques is accomplished by i) increasing the resolution of key signals in the NMR spectrum and ii) eliminating the splitting of the key signals by an adjacent NMR active nucleus. These effects are accomplished by specifically isotopically labeling at least one of the amino acids utilized in the synthesis of the protein with only those atoms that the analyst wishes to detect in the NMR spectrum, so that all other atoms, including those adjacent to the key nuclei, are unlabeled. This approach is a departure from current NMR labeling techniques wherein the goal has been to prepare proteins in a universally labeled form.

Proteins containing specifically labeled amino acids can be chemically synthesized or expressed by bacteria, yeast or mammalian or insect cells or in cell-free systems, as described by Yokoyama et al. The labeled proteins preferably comprise at least about 50 amino acid residues. The compositions and methods of the invention may advantageously be employed in connection with proteins having molecular masses of at least about 5kD.

If bacterial or yeast expression is desired, then the medium should contain all of the amino acids necessary for protein biosynthesis in the desired specifically labeled form to prevent non-specific labeling. Notwithstanding the provisions of

substantially all amino acids in specifically labeled form, isotope shuffling may still occur with bacteria or yeast grown in such a medium. Accordingly, proteins containing specifically isotopically labeled amino acids are preferably expressed either in a cell-free system or in mammalian or insect cells grown in a medium containing the amino acids required for protein biosynthesis. It is well known that nearly all naturally occurring amino acids cannot be synthesized by mammalian or insect cells, therefore, isotope shuffling will be at a minimum in such cells. The amino acid compositions for insect and mammalian cell culture media are well known. Such media are described in U.S. patents 5,393,669 and 5,627,044, the disclosures of which are incorporated herein by reference. Generally, all twenty essential amino acids are present in such media, and in accordance with the present invention, any or all such amino acids may be specifically isotopically labeled.

The labeled amino acids of the target protein are labeled at specific positions with any combination of the NMR isotopes  $^2\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ , such that only those atoms desired to be detectable in the spectrum are NMR active. It will be recognized by those skilled in the art that a key set of identifications required in elucidating protein structure by NMR is obtained from the assignment of signals from the backbone of the protein, *i.e.*, in the signals between the  $\alpha$ -carbon of a given amino acid and the amino protons of the same and adjacent residues in the protein sequence. Grzesiek, S. and Bax, A. J., *J. Magn. Reson.*, vol. 96:432-440

(1992). In the Grzesiek *et al.* experiment (the "HNCA experiment"), less than optimal sensitivity and resolution were achieved due to the influence of neighboring atoms whose presence is not essential for 5 background structural assignments, but which nevertheless were detected due to the universal labeling strategies employed. These complications are reduced by employing only specifically labeled amino acids in accordance with this invention.

10 In the instant invention, the amino acids of the target protein are advantageously labeled at the  $\alpha$ -amino group with  $^{15}\text{N}$  and at the C-carbonyl and the  $\alpha$ -carbon with  $^{13}\text{C}$ , while the side chains are left unlabeled. In this way, the signals from the 15 C-carbonyl and the  $\alpha$ -carbon are uncoupled from each other using conventional NMR techniques. Importantly, the signal from the  $\alpha$ -carbon is not split into two parts by the adjacent  $\beta$ -carbon atom when that carbon is in the inactive,  $^{12}\text{C}$  form. This approach contrasts with 20 the method described by Matsuo *et al.*, *J. Magn. Reson.*, 113, 91-96 (1996), which uses a selective radio-frequency field to decouple the  $\beta$ -carbon resonances. This method lacks generality, particularly, for serine residues, where the  $\alpha$ -carbon and the  $\beta$ -carbon 25 resonances are insufficiently resolved.

In a particularly preferred aspect of the invention, all of the amino acids of the target protein are not only labeled at the  $\alpha$ -amino group with  $^{15}\text{N}$  and at the C-carbonyl and the  $\alpha$ -carbon with  $^{13}\text{C}$ , but are 30 also deuterated at the  $\alpha$ -carbon, the side chains being left unlabeled. In this way, in addition to the above advantages, the linewidth of the signals from each  $\alpha$ -

carbon is significantly narrowed because the carbon nucleus is no longer efficiently relaxed by an attached proton. This decrease in linewidth significantly increases the resolution of the distinct signals from each amino acid residue (Grzesiek et al., *J. Am. Chem. Soc.*, 115, 4369-4370 (1993)).

In a further preferred aspect of this invention, all of the amino acids of the target protein are not only labeled at the  $\alpha$ -amino group with  $^{15}\text{N}$  and at the C-carbonyl and the  $\alpha$ -carbon with  $^{13}\text{C}$ , but are also partially protonated at the  $\alpha$ -carbon. This approach preserves the advantage of line-narrowing mentioned in the previous paragraph, as well as permits the application of experiments that involve protonation at the  $\alpha$ -carbon. These experiments include those described for determining long-range structure in macromolecules, which experiments exploit the presence of residual dipolar couplings between atoms such as  $^{13}\text{C}$  and  $^1\text{H}$  in dilute liquid crystalline solutions. (Tjandra and Bax, *Science* 278, 1111-1114 (1997)) The angular information derived from these experiments may be used for determining the structures of large proteins (>40kDa). The present invention thus may be used in connection with these experiments to restrict the dipolar coupling information to N-H and C $\alpha$ -H spin pairs, which greatly simplifies the relevant NMR spectra.

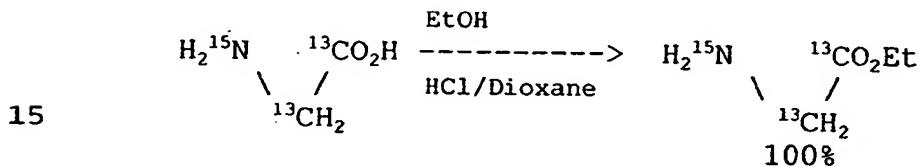
In this preferred aspect of the invention, the amino acids are deuterated at the  $\alpha$ -carbon to a level of about 30-70% in a preferred embodiment, about 40-60% in a more preferred embodiment, and about 50% in a most preferred embodiment.

Amino acids have been chemically synthesized in unlabeled forms by various means, and some have been synthesized in specifically isotopically labeled forms. See, e.g., Martin, *Isotopes Environ. Health Stud.*, 5 32:15 (1996); Schmidt, *Isotopes Environ. Health Stud.*, 31:161 (1995). Ragnarsson et al., *J. Chem. Soc. Perkin Trans 1*, 2503 (1994) synthesized BOC labeled forms of the following amino acids: 1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N Ala, Phe, Leu, and Tyr; 1,2-<sup>13</sup>C<sub>2</sub>, 3, 3, 3-<sup>2</sup>H<sub>3</sub>, <sup>15</sup>N Ala; 1,2-<sup>13</sup>C<sub>2</sub>, 3, 3-<sup>2</sup>H<sub>2</sub>, 10 <sup>15</sup>N Phe; 3, 3, 3-<sup>2</sup>H<sub>3</sub> Ala. Ragnarsson, *J. Chem. Soc. Chem. Commun.*, 935 (1996) also synthesized BOC labeled 1,2-<sup>13</sup>C<sub>2</sub>, 2-<sup>2</sup>H, <sup>15</sup>N Ala, Leu and Phe; and 1,2-<sup>13</sup>C<sub>2</sub>, 2,2-<sup>2</sup>H<sub>2</sub>, <sup>15</sup>N Gly which were partly used for conformational studies of the pentapeptide, Leu-Enkephalin 15 (*Biopolymers*, 41:591 (1997)). Unkefer (*J. Lab. Cpd. Radiopharm.*, 38:239 (1996)) synthesized <sup>15</sup>N labeled Ala, Val, Leu, and Phe as well as 1-<sup>13</sup>C, <sup>15</sup>N Val. However, as noted above, mammalian cell media require all twenty 20 amino acids for cell growth. In accordance with the present invention, methods for synthesizing all twenty amino acids in specifically labeled form and culture media containing all or any combination of such amino acids are provided.

Specifically isotopically labeled amino acids may 25 be synthesized by asymmetric synthesis from an appropriately isotopically labeled precursor. Glycine, specifically labeled with any combination of <sup>13</sup>C and <sup>15</sup>N, is readily available commercially. Preferably, therefore, the amino acids are synthesized using 30 glycine, isotopically labeled as required, as a precursor.

Methods for synthesizing amino acids from glycine have been described which may be used in accordance with the present invention (Duthaler, *Tetrahedron*, 50:1539 (1994); Schöllkopf, *Topics Curr. Chem.*, 109:65 (1983); Oppolzer, *Tet. Letts.*, 30:6009 (1989); *Helvetica Chimica Acta*, 77:2363 (1994); *Helvetica Chimica Acta*, 75:1965 (1992)).

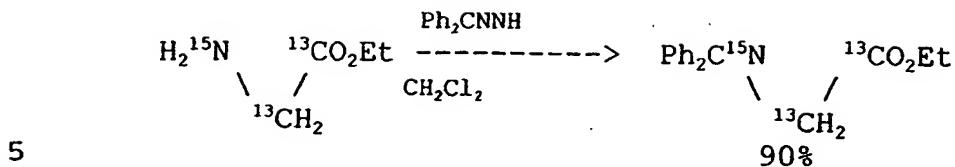
In one aspect of the invention,  $^{13}\text{C}_2$ ,  $^{15}\text{N}$ -glycine is first esterified with a suitable alcohol, such as methanol, ethanol or isopropanol, to give the corresponding ester.



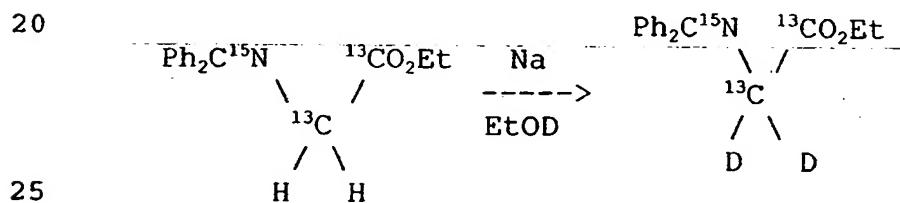
The amino group of the glycine ester may be protected by procedures known in the art. See Green, Protective Groups in Organic Synthesis, Wiley, N.Y. (1991).

20 Schiff bases (Stork, *J. Org. Chem.*, 41:3491 (1976)) are preferred for protection with the diphenyl ketimine (O'Donnell, *J. Org. Chem.*, 47:2663 (1982)) or bis(methylsulfanyl) imine (Hoppe, *Liebigs Ann. Chem.*, 1979, 2066) being particularly preferred. Introducing 25 the protecting group may be accomplished by reacting the glycine ester with the corresponding aryl imine for the diphenyl ketimine protecting group, or by reacting the glycine ester with carbon disulfide and methyl iodide for the bis(methylsulfanyl) imine protecting group.

30



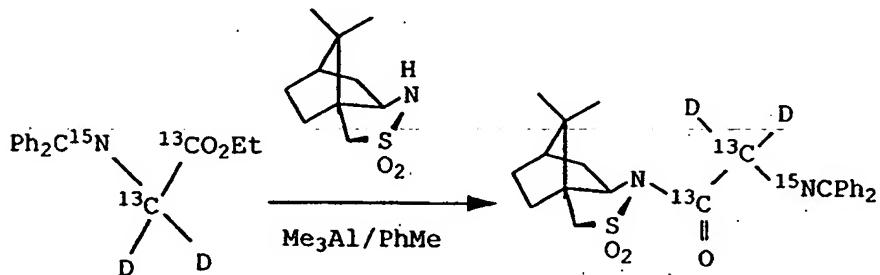
10 protected glycine derivative obtained above is deuterated at the  $\alpha$ -carbon by treating it with a base in a deuteronic solvent, such as sodium carbonate in  $D_2O$  (Ragnarsson, *J. Chem. Soc. Chem. Commun.*, 935 (1996)). To minimize loss of material due to hydrolysis of the ester function, the deuteration is preferably accomplished by treating the doubly protected glycine derivative with a catalytic amount of sodium in an anhydrous deuteronic solvent such as deuteromethanol (MeOD) or deuteroethanol (EtOD).



The required backbone labeled amino acids can now be synthesized from the doubly protected glycine derivative or, preferably, its deuterated analogue, by introducing the characteristic sidechain in a stereospecific manner to preserve the L-configuration at the  $\alpha$ -carbon chiral center. Methods for such chiral syntheses are known to those skilled in the art. They involve reacting the glycine derivative with a chiral

molecule, called a "chiral auxiliary," which directs the subsequent incorporation of the amino acid sidechain in a chiral manner (March, J., Advanced Organic Chemistry, 4th ed., Wiley, N.Y., p. 118, 1992).

5 In a particularly preferred aspect of the invention, the deuterated glycine analogue is converted to the chiral "sultam" derivative. See Oppolzer, J. *Chem. Soc. Perkin 1*: 2503 (1996). For example, methyl or ethyl N-[bis(methylthio)methylidene]glycinate or 10 methyl or ethyl N-(diphenyl methylene)glycinate is treated with (2R)-bornane-10,2-sultam or (2S)-bornane-10,2-sultam in the presence of trimethylaluminum or triethylaluminum and a solvent (usually toluene). (2R)-Bornane-10,2-sultam, ethyl N-(diphenyl 15 methylene)glycinate and trimethylaluminum are particularly preferred for forming the L-amino acids.



20 The resulting sultam derivative is then treated with a strong base such as lithium diisopropylamide ("LDA") or n-butyl lithium, in an appropriate solvent such as tetrahydrofuran ("THF"), in the presence of a coordinating solvent such as hexamethylphosphoramide ("HMPA") or N,N-dimethylpropyleneurea ("DMPU") to give the resulting glycine derivative.

To prepare amino acids with simple alkyl sidechains, i.e., alanine, leucine, isoleucine, phenylalanine, methionine, and valine, the derivatized glycine molecule is treated with the appropriate alkyl halide to form the fully protected amino acid. For example, treating the derivatized glycine molecule with benzyl iodide leads to the formation of protected phenylalanine. A list of alkyl halides and corresponding amino acids is provided in Table 1.

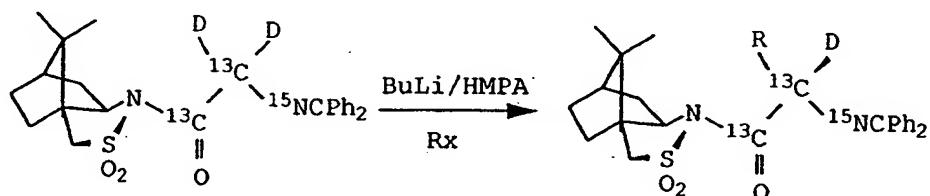
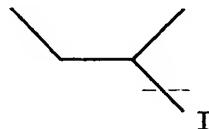


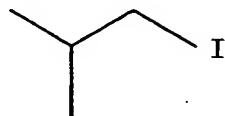
TABLE 1

### Alanine Me —

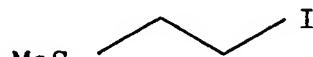
### Isoleucine



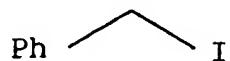
### Leucine



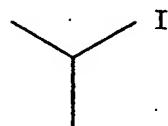
### Methionine



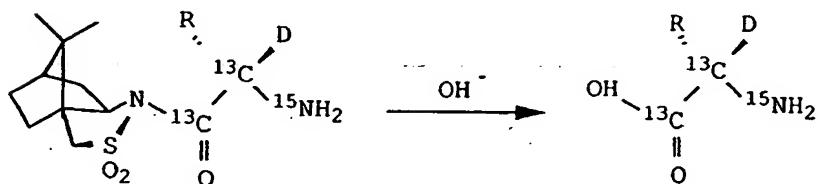
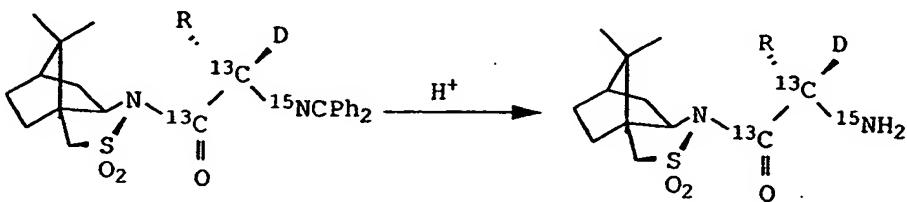
## Phenylalanine



### Valine



The fully protected amino acid thus prepared may be unblocked by a variety of means. The preferred method is a simple two-step procedure consisting of treating the protected amino acid with aqueous acid to remove the imine protecting group, followed by treating the amino acid with an aqueous base to remove the sultam group. In principle, any combination of an aqueous acid and base can be employed, but dilute HCl followed by dilute LiOH is preferred. The liberated, specifically isotopically labeled amino acid may then be further purified by, for instance, ion exchange chromatography.



To prepare aspartic acid, glutamic acid, tyrosine, histidine and tryptophan, the functional groups present in the sidechains are advantageously protected prior to reaction with the derivatized glycine molecule. Preferably, the derivatized glycine molecule is treated with a previously protected alkyl halide. For example, aspartic and glutamic acid may be prepared via the commercially available tert-butyl bromoacetate (Oppolzer, *Helvetica Chimica Acta*, 77:2363 (1994)) and

methyl acrylate (Schollkopf, *Synthesis*, 737 (1986)), respectively. The alkyl ester protecting group is removed by treating the glycine anion with acid during the two-step unblocking procedure described above to 5 give the desired amino acid.

Similarly, tyrosine may be prepared via the commercially available 4-benzyloxybenzyl or 4-methoxybenzyl chloride. The benzyl or methyl protecting group may be removed prior to the two-step 10 unblocking procedure by, for instance, treating the derivatized glycine molecule with trimethyl silyl iodide in a suitable solvent such as dichloromethane.

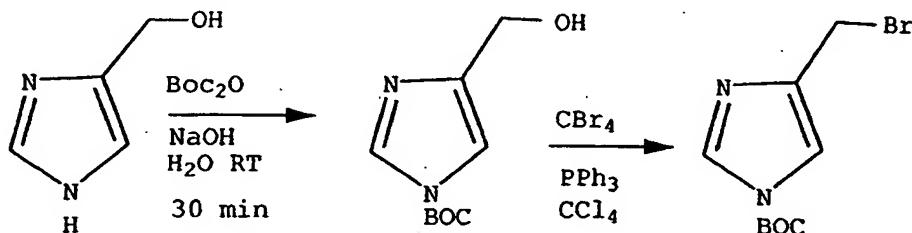
Protected sidechain precursors for histidine and tryptophan may be prepared, for example, by the 15 reaction shown in Table 2. For the preparation of the histidine precursor, commercially available 4-hydroxymethyl imidazole hydrochloride is protected at the ring amino nitrogen by a suitable protecting group such as t-boc, F-moc, tosyl, etc. The alcohol 20 functional group of the protected molecule is then converted to a suitable leaving group, e.g., the corresponding halide such as bromide, by reacting the alcohol with a suitable brominating agent, such as free bromine, or triphenylphosphine and carbon tetrabromide, 25 in a suitable solvent such as carbon tetrachloride. The protected bromomethylimidazole derivative may then be reacted directly with the derivatized glycine molecule.

Similarly, the required tryptophan precursor may 30 be prepared from commercially available indole-3-carboxaldehyde via protection of the ring nitrogen with

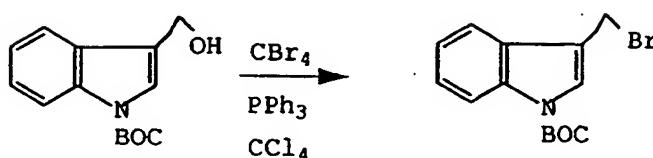
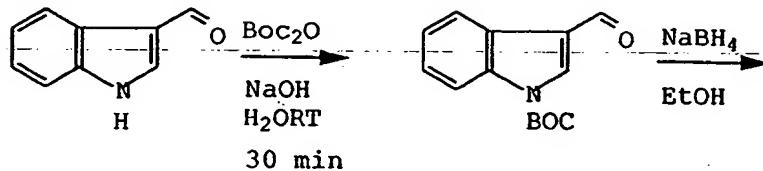
a suitable protecting group such as t-boc, F-moc, etc., followed by conversion to the corresponding alcohol by reduction with, for example, sodium borohydride in ethanol, and halogenation as described above. The 5 protected bromomethylindole derivative may then be reacted directly with the derivatized glycine molecule. The production of these heterocyclic halides and corresponding amino acids is illustrated in Table 2.

Table 2

## Histidine



## Tryptophan



10 Fully protected tryptophan and histidine may be unblocked by the simple two-step procedure described above as t-boc, F-moc, or tosyl groups may be removed by the acid/base treatment. Again, in principle any

combination of an aqueous acid or base can be employed. However, aqueous HCl followed by LiOH is preferred.

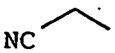
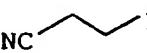
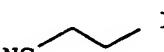
Specifically isotopically labeled asparagine and glutamine may be prepared respectively from labeled 5 aspartic acid and glutamic acid prepared above using established techniques. For example, the techniques described in U.S. patents 5,393,669 and 5,627,044 may be used. Alternatively, asparagine and glutamine, and arginine and lysine, can be prepared by treating the 10 derivatized glycine molecule with an alkyl halide carrying a terminal nitrile group. For example, treating the derivatized glycine molecule with 3-bromopropionitrile leads to the formation of the corresponding fully protected nitrile derivative.

15 Following unblocking by the two-step acid/base treatment described above, the resulting amino acid nitriles are converted to the desired amino acids. For example, lysine may be formed by reacting 4-bromobutyronitrile with the derivatized glycine

20 molecule and then reducing the resulting nitrile with a suitable reducing agent such as sodium borohydride and cobalt chloride.

A list of amino acids, corresponding halo-alkyl nitriles and methods for their conversion are provided 25 in Table 3.

TABLE 3

Asparagine	NC 	partial hydrolysis
Arginine	NC 	see below
Glutamine	NC 	partial hydrolysis
Lysine	NC 	reduction

Preferably, arginine is prepared from the nitrile isolated from the two-step unblocking procedure by reducing the nitrile with sodium borohydride and cobalt chloride, followed by treating the resulting ornithine with O-methylisourea tosylate. The O-methylisourea tosylate compound is prepared from urea treated with methyl tosylate in the presence of basic copper II carbonate, followed by treatment with sodium sulfhydride (Kurtz, *J. Biol. Chem.*, 180:1259 (1949)).

10 The remaining specifically isotopically labeled amino acids required for a specifically labeled mammalian or insect cell medium, i.e., serine, cysteine and threonine, may be prepared, for example, by the enzymatic procedures described in U.S. patents 15 5,393,669 and 5,627,044 and the references cited therein using  $^{13}\text{C}_2$ ,  $^{15}\text{N}$  glycine and/or  $^2\text{H}_2$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  glycine as a precursor.

20 The specifically isotopically labeled amino acids thus prepared may be incorporated into a mammalian or insect cell medium individually or in any combination so that the protein expressed by the cells growing in the medium may be specifically labeled at the amino acid residues of choice. The composition and use of such medium for bacterial, yeast, mammalian and insect 25 cell lines are well known. The compositions described in U.S. patent 5,324,658 and in U.S. patents 5,393,669 and 5,627,044 may advantageously be used for the media of this invention.

30 NMR analysis of the specifically labeled protein thus produced may be used to interpret NMR data from

the same protein separately obtained in universally labeled form and thereby expedite the determination of the structure of the protein. For instance, application of the HNCA experiment to a specifically labeled 5 protein will enable the maximum sensitivity and resolution to be obtained for the determination of the protein backbone resonance assignments. The  $\text{C}\alpha$  resonance for each amino acid residue will exhibit a correlation with the amide nitrogen atom of the same 10 residue via the one-bond  $\text{C}\alpha\text{-Ni}$  coupling, which is then transferred to the amide proton using another transfer via the one-bond  $\text{Ni}\text{-H}$  coupling. In addition, certain residues will exhibit a two-bond  $\text{C}\alpha\text{-1-}(\text{C}\text{i-1})\text{-Ni}$  15 correlation to the previous residue in such cases where this two-bond coupling is of sufficient magnitude. These latter data can be complemented by data from an experiment known as  $\text{H}\text{N}(\text{CO})\text{CA}$  which exhibits exclusively all such two-bond correlations due to transfer via the intervening carbonyl carbon. This latter experiment 20 also shares the advantages gained by the HNCA experiments with respect to selective labeling. Hence, the HNCA and  $\text{H}\text{N}(\text{CO})\text{CA}$  experiments combined, can be used sequentially to assign the backbone resonances of 25 proteins with high-sensitivity, and with sufficient resolution to permit automated analysis with computational algorithms.

The invention is illustrated by the following examples which are for illustrative purposes only and in no way limit the scope of the invention.

EXAMPLESExample 1

Synthesis of ethyl N-(diphenylmethylene) [1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2,2-<sup>2</sup>H<sub>2</sub>]glycinate.

5 Under anhydrous conditions, 4M HCL in dioxane (50 ml) was added to a solution of [1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N]glycine (5 g, 66.6 mmol) in ethanol (100 ml) and refluxed for 1 hour. Evaporating the solvent *in vacuo* and repeating the procedure twice yielded a white crystalline solid.

10 Benzophenone imine (1 eq.) was added to the ethyl glycinate hydrochloride in dry dichloromethane (100 ml) and stirred at room temperature overnight. Solid was filtered off and the solvent removed *in vacuo*. The resulting ethyl N-(diphenylmethylene) [1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N]glycinate was recrystallized from hexane.

15 Under anhydrous conditions, sodium metal (198 mg, 8.6 mmol) was added to a solution of ethyl N-(diphenylmethylene) [1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N]glycinate (19.31 g, 71.25 mmol) in freshly distilled deuteroethanol (250 ml, 60 eq). After stirring overnight at room temperature, the reaction was quenched by adding deuteroacetic acid (0.5 g, 8.6 mmol). Removing the solvent *in vacuo*, resuspending in dichloromethane (100 ml), filtering and evaporating yielded a white crystalline solid. Recrystallizing from hexane/ethyl acetate gave ethyl N-(diphenylmethylene) [1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2,2-<sup>2</sup>H<sub>2</sub>]glycinate (17 g, 87%).

Example 2

Synthesis of (2R)-N-(diphenylmethylene) [1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2,2-<sup>2</sup>H<sub>2</sub>]glycylbornane-10,2-sultam.

Over 20 minutes, 2 M trimethylaluminum in hexane (40 ml, 1.2 eq) was added to (2R)-bornane-10,2-sultam (15.7 g, 1.1 eq) in toluene (110 ml) at 0°C, then left for 30 minutes. Ethyl N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, 5 <sup>15</sup>N, 2,2-<sup>2</sup>H<sub>2</sub>]glycinate (18.1 g, 66.4 mmol) in toluene (10 ml) was added and the reaction stirred overnight. Heating to 50°C for 3-4 hours drove the reaction to completion. Workup was effected by cooling the reaction in ice and carefully adding MeOD (20 ml). 10 After 1 hour, D<sub>2</sub>O (30 ml) was carefully added. Filtering, extracting with ethyl acetate (2 x 250 ml), drying over MgSO<sub>4</sub>, and purifying by silica gel flash chromatography (Hexane: Ethyl Acetate 10:1 to 1:1) gave (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, 15N, 15 2,2-<sup>2</sup>H<sub>2</sub>]glycylbornane-10,2-sultam (28.7 g, 99% yield) as an orange oil.

Example 3

Synthesis of (1,2-<sup>13</sup>C<sub>2</sub>, 15N, 2-<sup>2</sup>H) Valine

Under anhydrous conditions, a solution of n-butyl lithium (2.5 M solution in hexane, 5.39 ml, 1.1 eq) was 20 added to a stirred solution of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, 15N, 2,2-<sup>2</sup>H<sub>2</sub>]glycylbornane-10,2-sultam (5.4 g, 12.3 mmol) in dry THF (120 ml) at -78°C. After 15 minutes, the 25 resulting solution was treated with HMPA (21.3 ml, 10 eq). After 1 hour, 2-iodopropane (6.12 ml, 5 eq) was added and the temperature raised to -10°C. After 2 days, the reaction was warmed to room temperature and quenched by adding D<sub>2</sub>O (50 ml). Extracting with diethyl 30 ether (100 ml), drying, evaporating and purifying by

silica gel chromatography (hexane:ethyl acetate 80:20) yielded 4.85 g (82%) of a semi-crystalline oil.

Deprotection was effected by adding 0.1 M HCL (100 ml) to a solution of the oil in THF (100 ml). After 15 minutes, lithium hydroxide (2.11 g, 5 eq) was added and the reaction stirred at room temperature overnight. Removing the solvent *in vacuo*, extracting with diethyl ether (5 x 50 ml) then with hexane (50 ml) and purifying the aqueous phase by ion exchange chromatography (Dowex 8 x 400 H<sup>+</sup> resin) gave the title compound (889 mg, 60%) as a white powder.

Example 4

Synthesis of (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2-<sup>2</sup>H) Phenylalanine

A solution of n-butyl lithium (2.5 M soln. in hexane, 10 ml, 1.6 eq) was added to a stirred solution of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2,2-<sup>2</sup>H<sub>2</sub>]glycylbornane-10,2-sultam (4.32 g, 9.8 mmol) in dry THF (50 ml) at -78°C. After 15 minutes, the resulting solution was treated with HMPA (13 ml, 7 eq) and the temperature raised to -50°C. After 1 hour, benzyl iodide (11.6 g, 5 eq) in THF (50 ml) was added. After 1 hour, the reaction was warmed to room temperature and quenched by adding water (50 ml). Extracting with diethyl ether (5 x 100 ml), washing with water, drying, and evaporating yielded an oil which was immediately deprotected.

Deprotection was effected by adding 0.2 M HCL (160 ml) to a solution of the oil in THF (160 ml). After 15 minutes, lithium hydroxide (6.71 g, 10 eq) was added and the reaction stirred at room temperature overnight. Removing the solvent *in vacuo*, extracting with diethyl

ether (5 x 50 ml) then with hexane (50 ml), and purifying the aqueous phase by ion exchange chromatography (Dowex 8 x 400 H<sup>+</sup> resin) gave the title compound (1.2 g, 45%) as a white powder.

5      Example 5

Synthesis of (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N) Alanine

A solution of n-butyl lithium (2.5 M soln in hexane, 2.8 ml, 1.1 eq) was added to a stirred solution of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N]glycylbornane-10,2-sultam (2.77 g, 6.3 mmol) in dry THF (63 ml) at -78°C. After 15 minutes, the resulting solution was treated with HMPA (11 ml, 10 eq). After 1 hour, methyl iodide (2 ml, 5 eq) was added, the reaction temperature raised to -10°C, stirred overnight, then quenched by adding water (20 ml). Extracting with diethyl ether (5 x 20 ml), washing with water, drying (MgSO<sub>4</sub>), evaporating and purifying by silica gel chromatography (hexane:ethyl acetate 80:20) yielded 2.34 g (82%) of a yellow crystalline solid.

20      Deprotection was effected by adding 0.2 M HCl (60 ml) to a solution of the crystals in THF (70 ml). After 15 minutes, lithium hydroxide (2.17 g, 10 eq) was added and the reaction stirred at room temperature overnight. Removing the solvent *in vacuo*, extracting with diethyl ether (5 x 50 ml) and purifying the aqueous phase by ion exchange chromatography (Dowex 8 x 400 H<sup>+</sup> resin) gave the title compound (395 mg, 68%) as a white powder.

Example 6

30      NMR Analysis of Specifically Labeled Phenylalanine

Backbone labeled (>95%  $^{13}\text{C}$ ,  $^{15}\text{N}$ , >90%  $^2\text{H}$ ) Phe (20 mg) (Figure 1(a)) was dissolved in 10 ml saturated  $\text{NaHCO}_3$ , to which was added 5 mole equivalents of acetic anhydride over a period of two hours. Following 5 desalting on a mixed-bed anion and cation exchange resin, the sample was prepared for NMR studies by dissolution in 700 ml  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (95:5 v/v). Two-dimensional HNCA spectra were acquired with deuterium decoupling on said acylated derivative of 10 backbone-labeled Phe versus the acylated derivative of uniformly triple (>95%  $^{13}\text{C}$ ,  $^{15}\text{N}$ , 50%  $^2\text{H}$ ) labeled Phe (Fig. 1(b)). As indicated, these spectra are shown in Figures 1(a) and 1(b), respectively.

Example 7

15 The acylated derivative of backbone ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ) labeled and uniformly  $^2\text{H}$  enriched Phe was dissolved in 700  $\mu\text{l}$  glycerol-d7/ $\text{H}_2\text{O}$  (85% v/v) to a final concentration of 1mM. A sample of the acylated derivative of uniformly ( $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^2\text{H}$ ) Phe was similarly 20 prepared. Two-dimensional HNCA spectra with  $^2\text{H}$  decoupling were acquired identically on these samples at 0°C. At this temperature, the rotational correlation time of the molecule (~18 ns) and hence resonance linewidths equaled that expected for a ~40 25 kDa protein. These spectra are shown in Figures 2(a) and 2(b), respectively.

Example 8

Synthesis of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2-<sup>1</sup>H, 2-<sup>2</sup>H]glycylbornane-10,2-sultam.

Over 20 minutes, 2 M trimethylaluminium in hexane (10.55 ml, 1.0 eq) was added to (2R)-bornane-10,2-sultam (6.11 g, 1.02 eq) in toluene (50 ml) at 0°C, then left for 30 minutes. Ethyl N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N]glycinate (6.11 g, 22.8 mmol) in toluene (17 ml) was added and the reaction stirred overnight. Heating to 60°C for 4-6 hours drove the reaction to completion. Workup was effected by cooling the reaction in ice and carefully adding MeOD (9.1 ml). After 1 hour, D<sub>2</sub>O (10.1 ml) was carefully added. Filtering, extracting with ethyl acetate (2 x 250 ml), and drying over NaSO<sub>4</sub> gave (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2-<sup>1</sup>H, 2-<sup>2</sup>H]glycylbornane-10,2-sultam (9.9 g, 96% yield) as an orange foam.

Example 9

Synthesis of (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 50% 2-<sup>2</sup>H) Valine

Under anhydrous conditions, a solution of n-butyl lithium (2.5 M soln in hexane, 5.5 ml, 1.1 eq) was added to a stirred solution of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2-<sup>1</sup>H, 2-<sup>2</sup>H]glycylbornane-10,2-sultam (5.49 g, 12.6 mmol) in dry THF (100 ml) at -78°C. After 15 minutes, the resulting solution was treated with HMPA (21.3 ml, 10 eq). After 1 hour, 2-iodopropane (6.12 ml, 5 eq) was added and the reaction was warmed to room temperature overnight and quenched by adding D<sub>2</sub>O (2.2 ml). Extraction with diethyl ether (100 ml), drying and

evaporation yielded 6.01 g (87.4%) of a white crystalline solid.

Deprotection (of 2.74 g, 5.74 mmol) was effected by adding 1 M HCL (9.06 ml) to a solution of the solid in THF (150 ml) and water (60 ml). After 15 minutes, lithium hydroxide (0.88 g, 3.65 eq) was added and the reaction stirred at room temperature overnight. Removing the solvent *in vacuo*, extracting with diethyl ether (5 x 50 ml), then with hexane (50 ml), and purifying the aqueous phase by ion exchange chromatography (Dowex 8 x 400 H<sup>+</sup> resin) gave the title compound (560 mg, 83%) as a white powder.

Example 10

Synthesis of (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 50% 2-<sup>2</sup>H) Phenylalanine

15 A solution of n-butyl lithium (2.5 M soln. in hexane, 5.77 ml, 1.1 eq) was added to a stirred solution of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2-<sup>1</sup>H, 2-<sup>2</sup>H]glycylbornane-10,2-sultam (5.77 g, 13.24 mmol) in dry THF (200 ml) at -78°C. After 15 minutes, 20 the resulting solution was treated with HMPA (22.7 ml, 10 eq). After 1 hour, benzyl bromide (7.87 ml, 5 eq) was added. The reaction was warmed to room temperature overnight and quenched by adding D<sub>2</sub>O (2.36 ml). Extracting with diethyl ether (5 x 100 ml), washing 25 with water, drying, and evaporating yielded an oil which was immediately deprotected.

Deprotection was effected by adding 1 M HCL (10.93 ml) to a solution of the oil in THF (72.8 ml) and water (72.8 ml). After 15 minutes, lithium hydroxide (1.07 g, 4.2 eq) was added and the reaction stirred at room temperature overnight. Removing the solvent *in vacuo*,

extracting with diethyl ether (5 x 50 ml), then with hexane (50 ml), and purifying the aqueous phase by ion exchange chromatography (Dowex 8 x 400 H<sup>+</sup> resin) gave the title compound (0.24 g, 23%) as a white powder.

5      Example 11

Synthesis of (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 50% 2-<sup>2</sup>H) Leucine

A solution of n-butyl lithium (2.5 M soln. in hexane, 5.77 ml, 1.1 eq) was added to a stirred solution of (2R)-N-(diphenylmethylene)[1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 2-<sup>1</sup>H, 2-<sup>2</sup>H]glycylbornane-10,2-sultam (5.77 g, 13.24 mmol) in dry THF (200 ml) at -78°C. After 15 minutes, the resulting solution was treated with HMPA (22.74 ml, 10 eq). After 1 hour, 1-iodo-2-methyl propane (12.18 g, 5 eq) was added. The reaction was warmed to room temperature overnight and quenched by adding D<sub>2</sub>O (2.36 ml). Extraction with diethyl ether (5 x 100 ml), washing with water, drying, and evaporation yielded a solid (6.49 g, 99.7%) which was immediately deprotected.

20      Deprotection was effected by adding 1 M HCl (12.1 ml) to a solution of the solid in THF (110 ml) and water (81 ml). After 15 minutes, lithium hydroxide (1.1 g, 3.5 eq) was added and the reaction stirred at room temperature overnight. Removing the solvent in vacuo, extracting with diethyl ether (5 x 50 ml), then with hexane (50 ml), and purifying the aqueous phase by ion exchange chromatography (Dowex 8 x 400 H<sup>+</sup> resin) gave the title compound (0.976 g, 97 %) as a white powder.

25

Example 12

To one liter of CHO S SFM, serum-free media (Life Technologies), supplied by the manufacturer with amino acids, pyruvate and carbohydrate omitted (Catalog No.

5 0920261) at 37°C were added 212 mg (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 50% 2-<sup>2</sup>H)leucine, 162 mg (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 50% 2-<sup>2</sup>H)valine, and 188 mg of backbone labeled (1,2-<sup>13</sup>C<sub>2</sub>, <sup>15</sup>N, 50% 2-<sup>2</sup>H)phenylalanine. The remaining unlabeled components were added as follows:

10 ten milliliters of sodium pyruvate 100x solution from Life Technologies (Catalog No. 11360-070),

3.9 grams of glucose,

20 mg aspartic acid,

31 mg glutamic acid,

15 57 mg asparagine,

82 mg histidine,

820 mg glutamine,

240 mg proline,

240 mg arginine,

20 135 mg threonine,

155 mg tyrosine,

60 mg methionine,

21 mg tryptophan,

210 mg isoleucine,

25 291 mg lysine,

18 mg alanine

17 mg glycine

48 mg serine

81 mg cysteine,

30 81 mg cystine, and

7.4 mg hydroxyproline.

The components were mixed for ten minutes, sonicated for three one-minute intervals, stirred for ten more minutes, and sterile filtered with a 0.2 m PES Nalgene sterile filter. The filtered mixture was 5 transferred to a Nalge bottle for shipping.

The resulting medium was used to culture a CHO cell line engineered to express human choriogonadotropin ("hCG"). Cells were cultured and the specifically isotopically labeled hCG  $\beta$ -subunit was 10 purified by procedures known in the art.

Example 13

Backbone labeled (Phe, Val, Leu) hCG  $\beta$ -subunit (~2.3 mg) was dissolved in 650 ul 100 mM phosphate buffer, and 50 ul 99.96% D<sub>2</sub>O was added for the field/frequency lock. Two-dimensional HNCA spectra without deuterium decoupling (Figure 3) and with deuterium decoupling (Figure 4) were acquired at 45°C with spectral widths of 3600 Hz and 1200 Hz in the <sup>1</sup>H and <sup>13</sup>C dimensions, respectively. Totals of 256 transients were acquired for each increment in the <sup>13</sup>C dimension, resulting in total acquisition times of 22 hours.

We claim:

1. A labeled protein in which the at least one of the amino acids of the protein is isotopically labeled in its backbone structure, but not its side chain, with any combination of  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^2\text{H}$ .
2. The labeled protein of claim 1, wherein the  $\alpha$ -amino nitrogen of said amino acid is  $^{15}\text{N}$  and the  $\alpha$ -carbon and carboxyl carbon are  $^{13}\text{C}$ .
3. The labeled protein of claim 2, wherein any hydrogen atom bonded to the  $\alpha$ -carbon is  $^2\text{H}$ .
4. The labeled protein of claim 2, wherein any hydrogen atom bonded to the  $\alpha$ -carbon is about 30-70%  $^2\text{H}$ .
5. The labeled protein of claim 2, wherein any hydrogen atom bonded to the  $\alpha$ -carbon is about 40-60%  $^2\text{H}$ .
6. The labeled protein of claim 2, wherein any hydrogen atom bonded to the  $\alpha$ -carbon is about 50%  $^2\text{H}$ .
7. The labeled protein of claim 1, wherein at least 10 of the amino acids of the protein are isotopically labeled in their backbone structures, but not in their side chains.
8. The labeled protein of claim 1, wherein substantially all of the amino acids of the protein are isotopically labeled in their backbone structures, but not in their side chains.
9. The labeled protein of claim 3, wherein at least 10 of the amino acids of the protein are isotopically labeled.
10. The labeled protein of claim 3, wherein substantially all of the amino acids of the protein are isotopically labeled.

11. The labeled protein of claim 1, which comprises at least 50 amino acid residues.

12. The labeled protein of claims 1, 3, 7, 9, or 10, which has a molecular mass of at least about 5kD.

13. A nutrient medium for the cultivation of bacterial, yeast, mammalian or insect cell cultures, which comprises all amino acids required for protein biosynthesis and assimilable sources of carbohydrate, 5 essential minerals and growth factors, wherein at least one of the amino acids of the medium is isotopically labeled in its backbone structure, but not its side chain, with any combination of <sup>13</sup>C, <sup>15</sup>N and <sup>2</sup>H.

14. The nutrient medium of claim 13, which is for cultivation of a mammalian or insect cell culture.

15. The nutrient medium of claim 14, wherein the  $\alpha$ -amino nitrogen of said amino acid is <sup>15</sup>N and the  $\alpha$ -carbon and carboxyl carbon are <sup>13</sup>C.

16. The nutrient medium of claim 15, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is <sup>2</sup>H.

17. The nutrient medium of claim 15, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 30-70% <sup>2</sup>H.

18. The nutrient medium of claim 15, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 40-60% <sup>2</sup>H.

19. The nutrient medium of claim 15, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 50% <sup>2</sup>H.

20. The nutrient medium of claim 14, wherein at least 10 of the amino acids of the nutrient medium are

isotopically labeled in their backbone structures, but not in their side chains.

21. The nutrient medium of claim 14, wherein substantially all of the amino acids of the nutrient medium are isotopically labeled in their backbone structures, but not in their side chains.

22. The nutrient medium of claim 16, wherein at least 10 of the amino acids of the nutrient medium are isotopically labeled.

23. The nutrient medium of claim 16, wherein substantially all of the amino acids of the nutrient medium are isotopically labeled.

24. The nutrient medium of claims 22 or 23, which contains all twenty amino acids.

25. A method for making a labeled protein which comprises cultivating, under protein-producing conditions, a bacterial, yeast, mammalian or insect cell culture capable of producing said protein on a nutrient medium which contains all amino acids required for protein biosynthesis and assimilable sources of carbohydrate, essential minerals and growth factors, wherein at least one of the amino acids of the medium is isotopically labeled in its backbone structure, but not its side chain, with any combination of  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^2\text{H}$ .

26. The method of claim 25, in which the cell culture is a mammalian or insect cell culture.

27. The method of claim 26, wherein the  $\alpha$ -amino nitrogen of said amino acid is  $^{15}\text{N}$  and the  $\alpha$ -carbon and carboxyl carbon are  $^{13}\text{C}$ .

28. The method of claim 27, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is  $^2\text{H}$ .

29. The method of claim 27, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 30-70%  $^2\text{H}$ .

30. The method of claim 27, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 40-60%  $^2\text{H}$ .

31. The method of claim 27, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 50%  $^2\text{H}$ .

32. The method of claim 26, wherein at least 10 of the amino acids of the nutrient medium are isotopically labeled in their backbone structures, but not in their side chains.

33. The method of claim 26, wherein substantially all of the amino acids of the nutrient medium are isotopically labeled in their backbone structures, but not in their side chains.

34. The method of claim 28, wherein at least 10 of the amino acids of the nutrient medium are isotopically labeled.

35. The method of claim 28, wherein substantially all of the amino acids of the nutrient medium are isotopically labeled.

36. The method of claims 34 or 35, wherein the nutrient medium contains all twenty amino acids.

37. A method for determining three-dimensional structure information of a protein, comprising the steps of:

5 (a) producing said protein in isotopically labeled form by cultivating, under protein-producing conditions, a bacterial, yeast, mammalian or insect cell culture that is capable of producing the protein

in a nutrient medium containing all amino acids required for protein biosynthesis and containing assimilable sources of carbohydrate, essential minerals and growth factors, wherein at least one of the amino acids of the medium is isotopically labeled in its backbone structure, but not its side chain, with any combination of  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^2\text{H}$ .

5 (b) isolating the labeled protein from the nutrient medium; and

10 (c) subjecting the protein to NMR spectroscopic analysis to determine information about its three-dimensional structure.

38. The method of claim 37, wherein the isotopically labeled protein is produced in a mammalian or insect cell culture.

39. The method of claim 37, wherein the  $\alpha$ -amino nitrogen of said amino acid is  $^{15}\text{N}$  and the  $\alpha$ -carbon and carboxyl carbon are  $^{13}\text{C}$ .

40. The method of claim 39, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is  $^2\text{H}$ .

41. The method of claim 39, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 30-70%  $^2\text{H}$ .

42. The method of claim 39, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 40-60%  $^2\text{H}$ .

43. The method of claim 39, wherein any hydrogen atom bonded to the  $\alpha$ -carbon of said amino acid is about 50%  $^2\text{H}$ .

44. The method of claim 37, wherein at least 10 of the amino acids of the nutrient medium are

isotopically labeled in their backbone structures, but not in their side chains.

45. The method of claim 37, wherein substantially all of the amino acids of the nutrient medium are isotopically labeled in their backbone structures, but not in their side chains.

46. The method of claim 40, wherein at least 10 of the amino acids of the nutrient medium are isotopically labeled.

47. The method of claim 40, wherein substantially all of the amino acids of the nutrient medium are isotopically labeled.

48. The method of claims 46 or 47, wherein the nutrient medium contains all twenty amino acids.

49. The method of claim 40, wherein the protein comprises at least 50 amino acid residues.

50. The method of claims 38, 39, 40, 46 or 47, wherein the protein has a molecular mass of at least about 5kD.

1/5

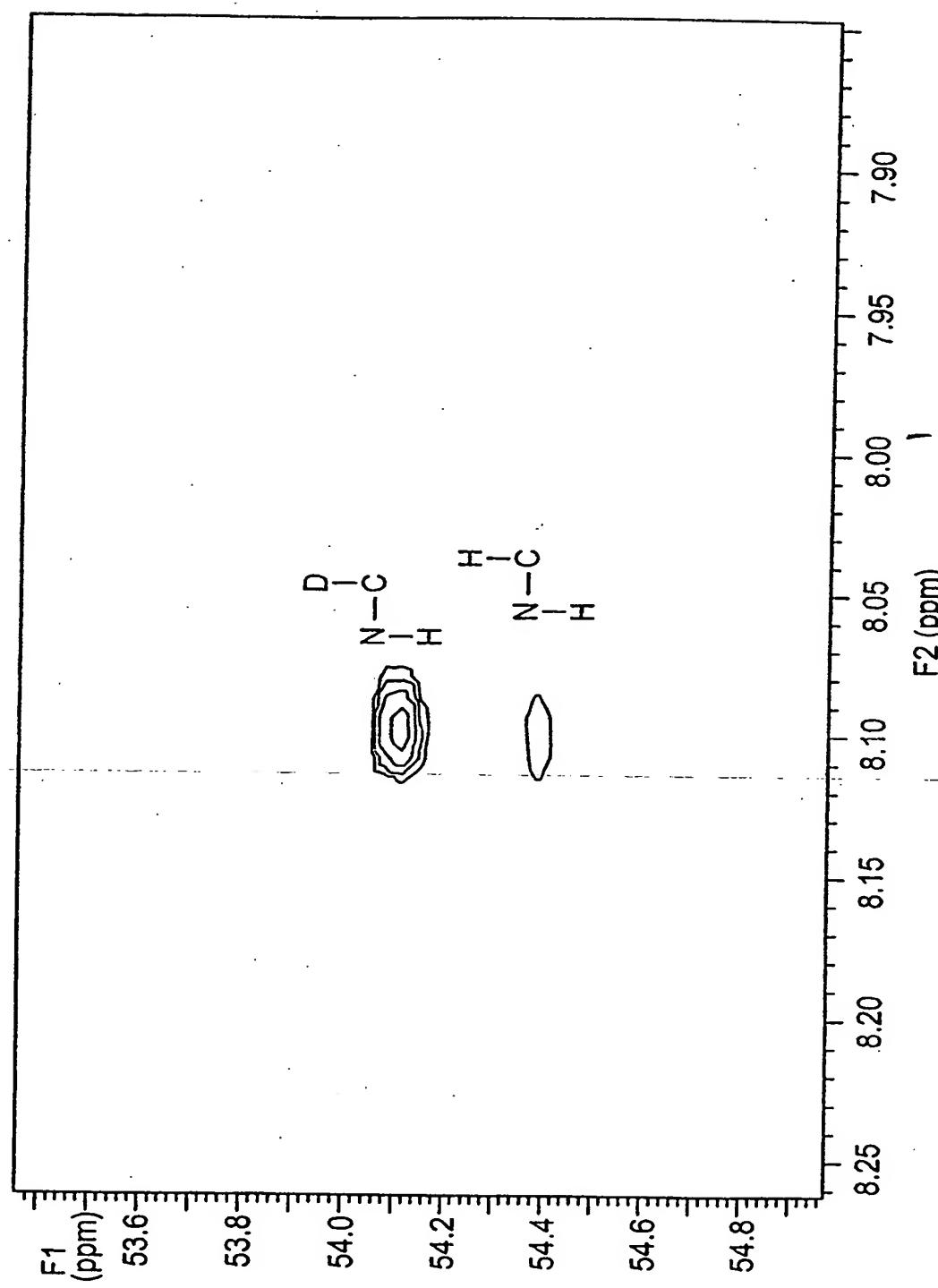


FIG. 1A

2/5

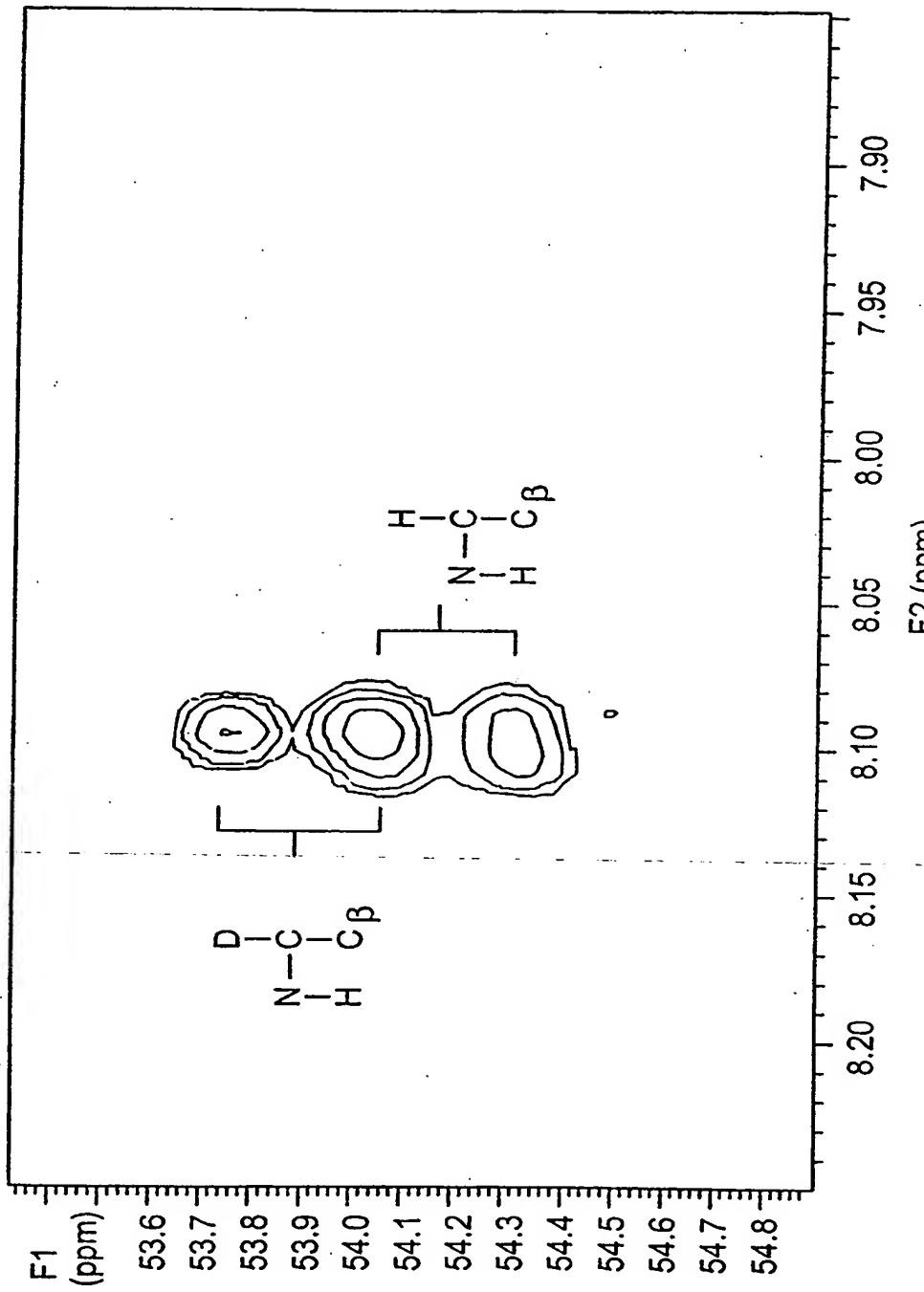
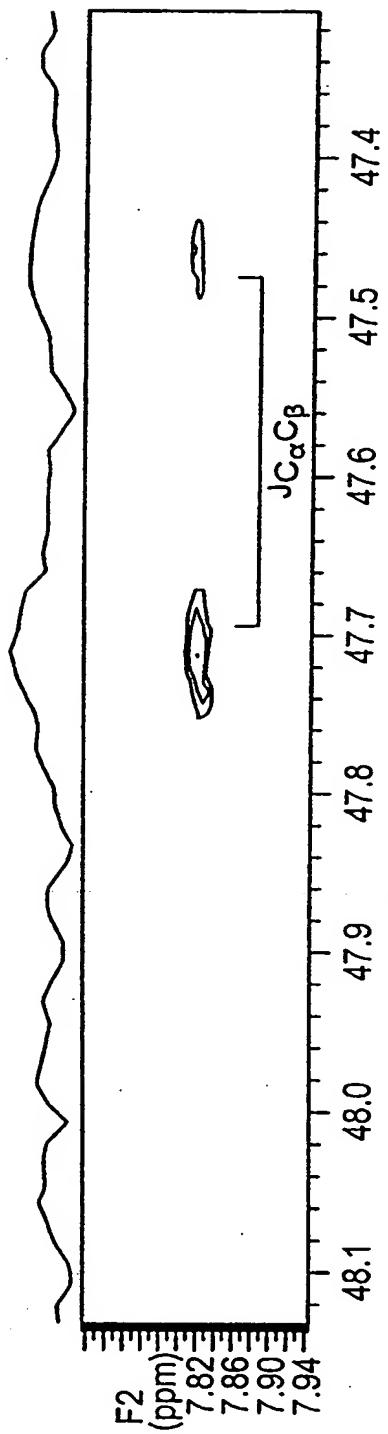
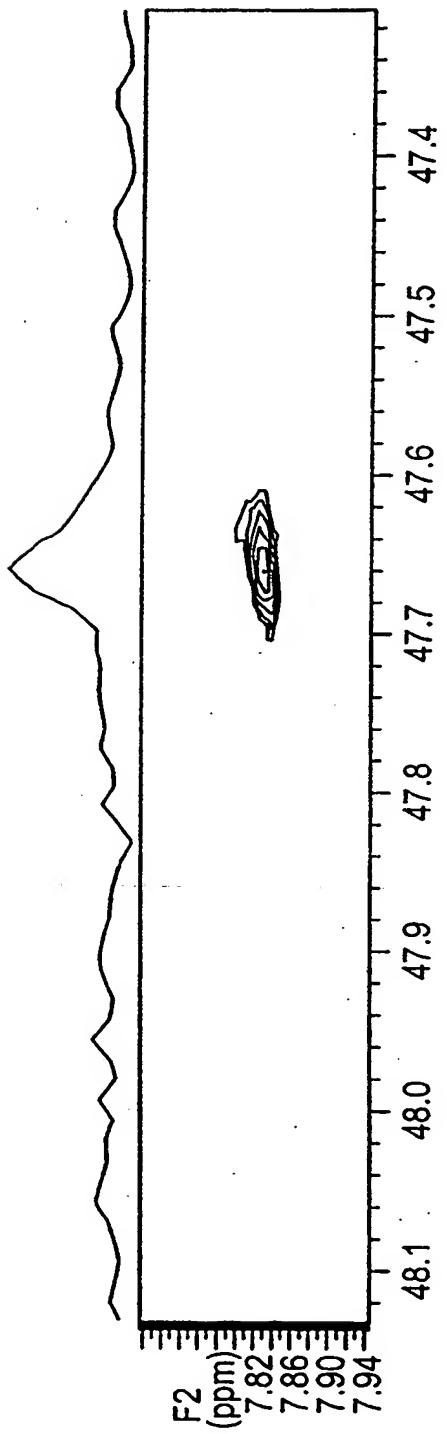


FIG. 1B

3/5



4/5

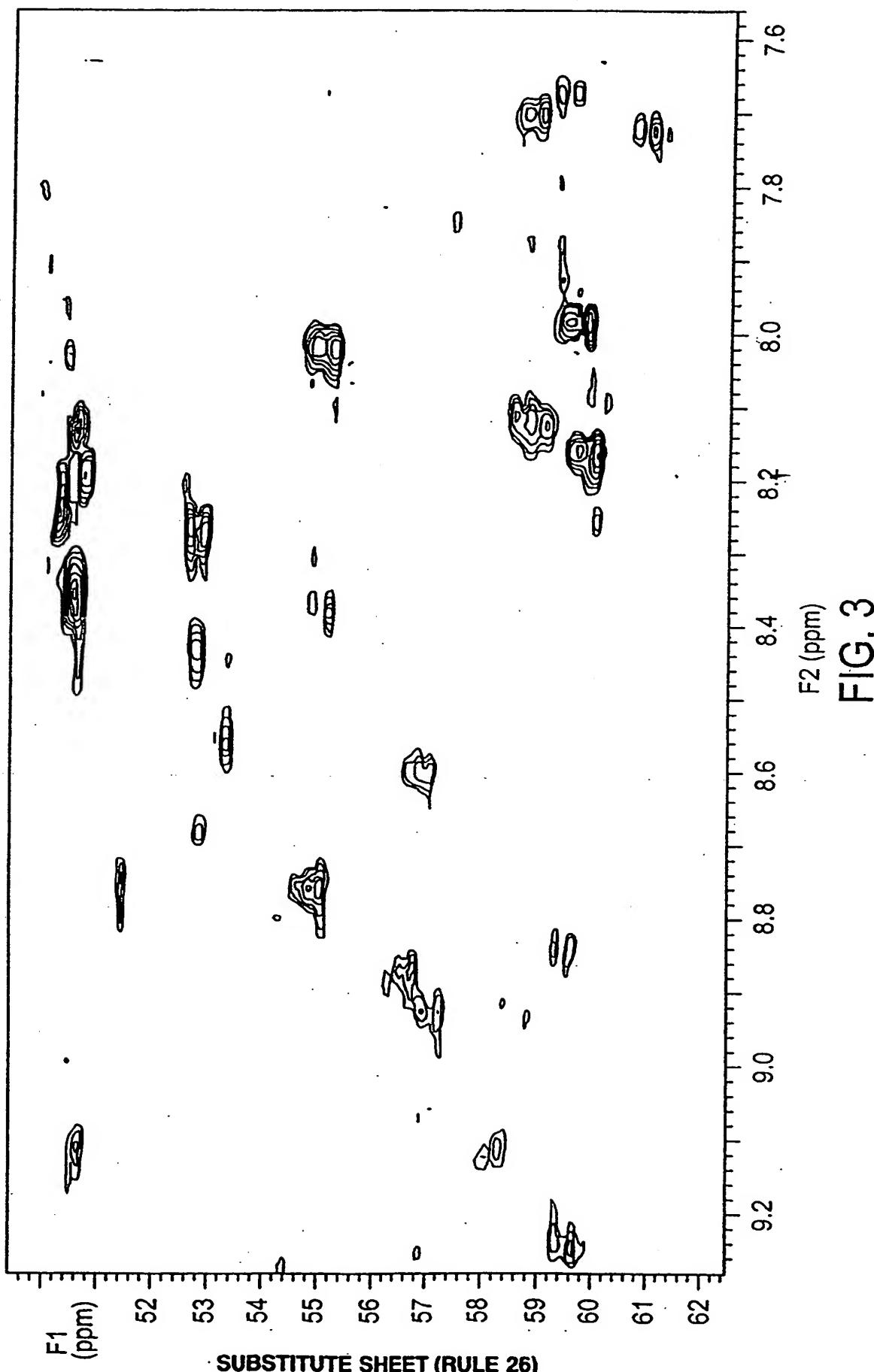


FIG. 3

5/5

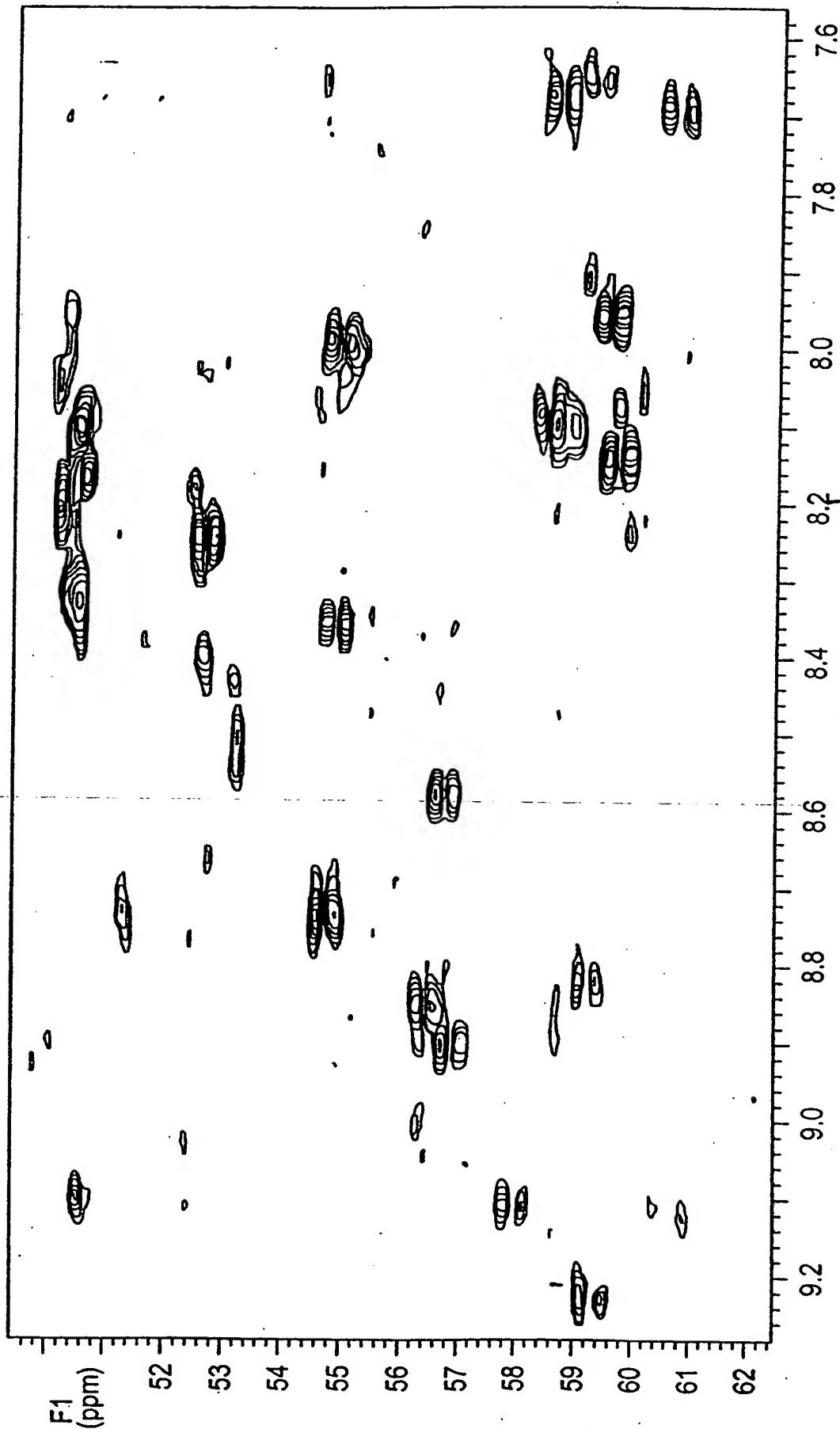


FIG. 4

# INTERNATIONAL SEARCH REPORT

Inte onal Application No  
PCT/US 98/18197

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C07B59/00 C12N1/20 C12N5/00 C12N5/06 C12P21/00  
G01N24/08 C07K14/59

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C07B C12N C12P G01N A61K C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NYASSÉ E.A.: "First synthesis of a fully '15N,13C!backbone-labelled peptide. 15N NMR spectrum of corresponding Leu-enkephalin" J.CHEM.SOC.,CHEMICAL COMMUNICATIONS.,1994, pages 2005-2006, XP002088044 see the whole document	1,2,8,10
Y	ARCHER E.A.: "TGF B1: NMR signal assignments of the recombinant protein expressed and isotopically enriched using Chinese hamster ovary cells" BIOCHEMISTRY, vol. 32, no. 4, 1993, pages 1152-1163, XP002088047 EASTON, PA US	1-50
X	ARCHE E.A.: "TGF B1: NMR signal assignments of the recombinant protein expressed and isotopically enriched using Chinese hamster ovary cells" BIOCHEMISTRY, vol. 32, no. 4, 1993, pages 1152-1163, XP002088047 EASTON, PA US	1,7, 11-14, 25,26, 37,38,50
Y	The whole document; see especially Table 1	1-50
	—	—/—

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the International search report

4 January 1999

15/01/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Groenendijk, M

## INTERNATIONAL SEARCH REPORT

Int'l Application No

PCT/US 98/18197

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 94 18339 A (MARTEK BIOSCIENCES CORP ; BROWN JONATHAN MILES (US)) 18 August 1994 cited in the application see the whole document ----	1-50
Y	LANKIEWICZ E.A.: "Synthesis of amino acid derivatives substituted in the backbone with stable isotopes for application in peptide synthesis" JOURNAL OF THE CHEMICAL SOCIETY, PERKIN TRANS I, vol. 17, 1994, pages 2503-2510, XP002088045 cited in the application see the whole document ----	1-50
Y	RAGNARSSON E.A.: "Proteinogenic amino acids labelled with <sup>15</sup> N and/or <sup>13</sup> C for application in peptide synthesis; a short review..." J. PEPTIDE SCIENCE, vol. 1, no. 3, 1995, pages 149-156, XP002089011 see the whole document ----	1-50

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Int'l	Application No
PCT/US 98/18197	

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9418339	A 18-08-1994	US 5393669 A		28-02-1995
		AU 3608797 A		04-12-1997
		AU 678446 B		29-05-1997
		AU 6236394 A		29-08-1994
		CA 2129904 A		18-08-1994
		EP 0636183 A		01-02-1995
		FI 943745 A		15-08-1994
		JP 7506192 T		06-07-1995
		NO 943703 A		23-11-1994
		US 5627044 A		06-05-1997
		US 5817474 A		06-10-1998